



UTILITY PATENT APPLICATION TRANSMITTAL

(Only for new nonprovisional applications under 37 CFR 1.53(b))

Attorney Docket No. 169,1327 CIP

First Named Inventor or Application Identifier

EDWIN HO

Express Mail Label No.

APPLICATION ELEMENTS

See MPEP chapter 600 concerning utility patent application contents.

ADDRESS TO:

Assistant Commissioner for Patents
Box Patent Application
Washington, DC 20231

1. ☐ Fee Transmittal Form
(Submit an original, and a duplicate for fee processing)
2. ☒ Specification Total Pages 33
3. ☒ Drawing(s) (35 USC 113) Total Sheets 6
4. ☒ Oath or Declaration Total Pages 2
- a. ☐ Newly executed (original or copy)
- b. ☒ Unexecuted for information purposes
- c. ☐ Copy from a prior application (37 CFR 1.63(d))
(for continuation/divisional with Box 17 completed)
[Note Box 5 below]
- i. ☐ **DELETION OF INVENTOR(S)**
Signed Statement attached deleting
inventor(s) named in the prior application, see
37 CFR 1.63(d)(2) and 1.33(b)
5. ☐ Incorporation By Reference (useable if Box 4c is checked)
The entire disclosure of the prior application, from which a copy of
the oath or declaration is supplied under Box 4c, is considered as
being part of the disclosure of the accompanying application and is
hereby incorporated by reference therein.

6. ☐ Microfiche Computer Program (Appendix)
7. Nucleotide and/or Amino Acid Sequence Submission
(if applicable, all necessary)
- a. ☐ Computer Readable Copy
- b. ☐ Paper Copy (identical to computer copy)
- c. ☐ Statement verifying identity of above copies

ACCOMPANYING APPLICATION PARTS

8. ☐ Assignment Papers (cover sheet & document(s))
9. ☐ 37 CFR 3.73(b) Statement ☐ Power of Attorney
(when there is an assignee)
10. ☐ English Translation Document (if applicable)
11. ☐ Information Disclosure Statement (IDS)/PTO-1449 ☐ Copies of IDS Citations
12. ☐ Preliminary Amendment
13. ☒ Return Receipt Postcard (MPEP 503)
(Should be specifically itemized)
14. ☐ Small Entity Statement(s) ☐ Statement filed in prior application
Status still proper and desired
15. ☐ Certified Copy of Priority Document(s)
(if foreign priority is claimed)
16. ☐ Other: _____

17. If a CONTINUING APPLICATION, check appropriate box and supply the requisite information:

- ☐ Continuation ☐ Divisional ☒ Continuation-in-part (CIP) of prior application No. 09/326,561 filed June 7, 1999

18. CORRESPONDENCE ADDRESS

☒ Customer Number or Bar Code Label

05514

(Insert Customer No. or Attach bar code label here)

or ☐ Correspondence address below

NAME

Address

City

State

Zip Code

Country

Telephone

Fax



CLAIMS	(1) FOR	(2) NUMBER FILED	(3) NUMBER EXTRA	(4) RATE	(5) CALCULATIONS
	TOTAL CLAIMS (37 CFR 1.16(c))	70-20 =	50	X \$ 18.00 =	\$ 900.00
	INDEPENDENT CLAIMS (37 CFR 1.16(b))	6-3 =	3	X \$ 78.00 =	\$ 234.00
	MULTIPLE DEPENDENT CLAIMS (if applicable) (37 CFR 1.16(d))			\$260.00 =	\$ 260.00
				BASIC FEE (37 CFR 1.16(a))	\$ 760.00
			Total of above Calculations =		\$2154.00
	Reduction by 50% for filing by small entity (Note 37 CFR 1.9, 1.27, 1.28).				
	TOTAL =				\$2154.00

19. Small entity status

- a. ☐ A Small entity statement is enclosed
- b. ☐ A small entity statement was filed in the prior nonprovisional application and such status is still proper and desired.
- c. ☐ Is no longer claimed.

20. ☒ A check in the amount of \$ \$2154.00 to cover the filing fee is enclosed.

21. ☐ A check in the amount of \$ _____ to cover the recordal fee is enclosed.

22. The Commissioner is hereby authorized to credit overpayments or charge the following fees to Deposit Account No. 06-1205:

- a. ☒ Fees required under 37 CFR 1.16.
- b. ☒ Fees required under 37 CFR 1.17.
- c. ☐ Fees required under 37 CFR 1.18.

SIGNATURE OF APPLICANT, ATTORNEY, OR AGENT REQUIRED

NAME

Leonard P. Diana (Reg. No. 29,296)

SIGNATURE

DATE

December 7, 1999

/PAP

NY_MAIN 46129 v 1

FACE DETECTION IN DIGITAL IMAGES

Technical Field of the Invention

The present invention relates to digital colour images and, in particular, to the detection of faces in colour digital images.

5

Background Art

Colour digital images are increasingly being stored in multi-media databases, and utilised in various computer applications. In many such applications it is desirable to be able to detect the location of a face in a visual image as one step in a multi-step process. The multi-step process can include content-based image retrieval, personal identification or verification for use with automatic teller machines or security cameras, or automated interaction between humans and computational devices.

15 Various prior art face detection methods are known including eigenfaces, neural networks, clustering, feature identification and skin colour techniques. Each of these techniques has its strengths and weaknesses, however, one feature which they have in common is that they are computationally intensive and therefore very slow, or they are fast but not sufficiently robust to detect faces.

20 The eigenface or eigenvector method is particularly suitable for face recognition and there is some tolerance for lighting variation, however it does not cope with different viewpoints of faces and does not handle occlusion of various facial features (such as occurs if a person is wearing sunglasses). Also it is not scale invariant.

25 The neural network approach utilises training based on a large number of face images and non-face images and has the advantages of being relatively simple to implement, providing some tolerance to the occlusion of facial features and some tolerance to lighting variation. It is also relatively easy to improve the detection rate by re-training the neural network using false detections. However, it is not scale invariant,

does not cope with different viewpoints or orientation, and leads to an exhaustive process to locate faces on an image.

5 The clustering technique is somewhat similar to the eigenface approach. A pixel window (eg. 20×20) is typically moved over the image and the distance between the resulting test pattern and a prototype face image and a prototype non-face image is represented by a vector. The vector captures the similarity and differences between the test pattern and the face model. A neural network can then be trained to classify as to whether the vector represents a face or a non-face. While this method is robust, it does not cope with different scales, different viewpoints or orientations. It leads to an
10 exhaustive approach to locate faces and relies upon assumed parameters.

The feature identification method is based upon searching for potential facial features or groups of facial features such as eyebrows, eyes, nose and mouth. The detection process involves identifying facial features and grouping these features into feature pairs, partial face groups, or face candidates. This process is advantageous in that
15 it is relatively scale invariant, there is no exhaustive searching, it is able to handle the occlusion of some facial features and it is also able to handle different viewpoints and orientations. Its main disadvantages are that there are potentially many false detections and that its performance is very dependent upon the facial feature detection algorithms used.

20 The use of skin colour to detect human faces is described in a paper by Yang J and Waibel A (1995) "Tracking Human Faces in Real-Time" CMU-CS-95-210, School of Computer Science, Carnegie Mellon University. This proposal was based on the concept that the human visual system adapts to different levels of brightness and to different illumination sources which implies that the human perception of colour is consistent
25 within a wide range of environmental lighting conditions. It was therefore thought

possible to remove brightness from the skin colour representation while preserving accurate, but low dimensional, colour information. As a consequence, in this prior art technique, the chromatic colour space was used. Chromatic colours (eg. r and g) can be derived from the RGB values as:

5.
$$r = R / (R + G + B) \quad \text{and} \quad g = G / (R + G + B)$$

These chromatic colours are known as "pure" colours in the absence of brightness.

Utilising this colour space, Yang and Waibel found the distribution of skin colour of different people, including both different persons and different races, was clustered together. This means that the skin colours of different people are very close and that the main differences are in differences of intensity.

This prior art method first of all generated a skin colour distribution model using a set of example face images from which skin colour regions were manually selected. Then the test image was converted to the chromatic colour space. Next each image in the test image (as converted) was then compared to the distribution of the skin colour model. Finally, all skin colour pixels so detected were identified, and regions of adjacent skin colour pixels could then be considered potential face candidates.

This prior art method has the advantage that processing colour is much faster than processing individual facial features, that colour is substantially orientation invariant and that it is insensitive to the occlusions of some facial features. The system is also substantially viewpoint invariant and scale invariant. However, the method suffers from a number of disadvantages including that the colour representation of the face can be influenced by different lighting conditions, and that different cameras (eg. digital or film) can produce different colour values even for the same person in the same environment.

However a significant disadvantage of the prior art methods is that the skin colour model is not very discriminating (ie. selecting pixels on a basis of whether they are included in the skin colour distribution results in a lot of non-skin colour pixels being included erroneously). It is also difficult to locate clusters or regions of skin colour pixels

5 that can be considered as candidate faces.

Disclosure of the Invention

An object of the present invention is to provide an improved method of detecting one or more faces in digital colour images.

In accordance with a first aspect of the present invention there is disclosed a

10 method of detecting a face in a colour digital image formed of a plurality of pixels, said method comprising the steps of:

- (i) testing the colour of said pixels to determine those said pixels having predominantly skin colour, said testing utilising at least one image capture condition provided with said image; and
- 15 (ii) subjecting only said those pixels determined in step (i) as having predominantly skin colour to further facial feature analysis whereby those said pixels not having a predominantly skin colour are not subjected to said further facial feature analysis.

Preferably each image capture condition is acquired at a time the image is

20 captured. Advantageously, the image is encoded according to a predetermined format and at least one image capture condition is represented as meta-data associated with the format. Most preferably the at least one image capture condition comprises lighting conditions at a time the image was captured.

In a particular implementation step (i), comprises the sub-step, preceding said

25 testing, of:

(a) dividing said image into a plurality of regions, each said region comprising a plurality of said pixels; and

wherein said testing is performed on pixels within each said region to determine those ones of said regions that are predominantly skin colour and step (ii) comprises
5 performing said further facial feature analysis on only those said regions determined to be predominantly of skin colour.

In accordance with another aspect of the present invention, there is disclosed a method of detecting a face in a colour digital image, said method comprising the steps of.

(i) segmenting said image into a plurality of regions each having a
10 substantially homogenous colour;

(ii) testing the colour of each said region created in step (i) to determine those regions having predominantly skin colour; and

(iii) subjecting only the regions determined in step (ii) to further facial feature analysis whereby said regions created in step (i) not having a predominantly skin
15 colour are not subjected to said further feature analysis.

Apparatus and computer readable media for performing the invention are also disclosed

Brief Description of the Drawings

A number of embodiments of the present invention will now be described with
20 reference to the drawings in which:

Fig. 1 is a schematic representation of the pixels of a colour digital image;

Fig. 2 shows the segmenting of the image of Fig. 1 into a plurality of regions each having a substantially homogenous colour according to a first embodiment;

Fig. 3¹ is a flow chart of a face detection process according to the first
25 embodiment;

Fig. 4 is a schematic block diagram of a general purpose computer upon which embodiments of the present invention can be practised;

Fig. 5 is a flow chart depicting the generation of a face colour distribution model;

Fig. 6 is a flow chart of a face detection process according to a second embodiment; and

Fig. 7 is a flow chart of a face detection process according to a third embodiment.

Detailed Description including Best Mode

Fig. 1 illustrates a typical colour digital image 1 having a size of 832 x 624 pixels 5, each of which has an RGB value.

According to a first embodiment of the present invention, rather than consider the skin colour of the image on a pixel by pixel basis as described above in relation to the prior art of Yang and Waibel, the image 1 is segmented into a number of regions. An example of such segmentation is schematically illustrated in Fig. 2 on the segmentation basis that all the pixels in each region 2 have substantially the same colour. Alternatively

the image may be segmented into arbitrary regions for processing.

The first embodiment implements a process 30 illustrated in the flow chart of Fig. 3, in which the regional segmentation of the image is carried out at step 31. Next, the regions of the image are converted in step 32 into the chromatic colour space (as described above). The next step 33 is to select those of the regions determined in step 31 which have a specified percentage (typically in the range 90-95%) of pixels having a skin colour. These selected regions are then conveniently represented by a boundary box or other boundary indication. Finally, at step 34 the selected regions, including any combinations of overlapping regions, are subjected to a further analysis (preferably not based on skin colour) to determine if the selected region(s) represent a face or faces.

This initial colour grouping can use any region based colour image segmentation technique. Preferably the image is partitioned into colour regions by seeking connected groups of pixels which have similar colours over a local region. Very small isolated initial spatial groupings can be ignored in order to find major colour regions and reduce a noise effect. A representative colour of each initial spatial region is determined by an average colour value of the region.

A colour region starts from an arbitrarily chosen pixel, which is compared with its neighbouring pixels. The region size is increased by adding neighbouring pixels, which are similar in colour, using a colour similarity threshold T . A neighbouring pixel is added to the region if $|R_p - R_m| < T$ and $|G_p - G_m| < T$ and $|B_p - B_m| < T$, where R_p , G_p , B_p are R , G , B values of the neighbouring pixel and R_m , G_m , B_m represented average R , G , B values of the region.

When a region has no more neighbouring pixels of similar colour, the region stops growing and represents one of the initial spatial groupings. If the region size is below a predetermined threshold value, it is ignored. A region having a pixel number equal or greater than the predetermined threshold is represented by its average colour.

A new pixel which does not yet belong to any region is chosen to start a new colour region. The process continues until each pixel in the image either belongs to an initial spatial grouping or has been ignored as being part of a small region.

The initial spatial groupings provide a colour region segmentation of the image with each region being represented by its average colour.

In this way, for most images where the bulk of the image is not a face, or part of a face, the majority of pixels will be grouped into regions or objects (be they foreground or background, etc) which are clearly not faces. Therefore these non-facial objects can be quickly eliminated on the basis of their colour.

- Once the regions have been determined, they are then converted into the "pure" chromatic colour space utilising the equations given above so as to provide r and g values. A generous rule such as a rule that at least 85% of the pixels within a given region be of face colour can be used to select those regions worthy of further examination. Preferably,
- 5 the test for face colour takes into account the nature of the original image, for example whether the image was taken with or without a flash. This information can be determined from the image source, eg. a camera.

- Thereafter, only those selected regions are subjected to a further test to determine the presence of facial features. This further test provides a conclusive determination as to
- 10 whether or not a region constitutes a face. In this connection, the further test is likely to be computationally slower and therefore the above described elimination of regions ensures that the computationally slow method is only applied to relatively small portions of the overall image. Thus the total processing time is reduced. Accordingly, the above method performs a computationally simple process on most, if not all pixels, and then
- 15 only performs complex examination on skin colour regions.

The preferred method of verifying if a region represents a face relies upon edge detection techniques as a means of detecting facial features. In particular facial features such as eyes, eyebrows and mouths often appear as dark bars on a face and thus provide dark edges.

- 20 The preferred form of edge detection is use of an edge detection filter. This utilises two functions operating in orthogonal directions. To detect a horizontal bar a second derivative Gaussian function is used in the vertical direction and a Gaussian function is used in the horizontal direction.

- Once an edge has been determined in this way each detected edge is examined.
- 25 Any pair of detected edges can be found to be derived from, and thus be indicative of, a

pair of eyes, a pair of eyebrows, or an eye and associated eyebrow, depending upon the relative position and size of the detected edges. Similarly, an individual edge can be derived from, and thus be indicative of, a mouth if it is located at an appropriate position relative to the eyes and/or eyebrows already detected.

5 By proceeding in this fashion, a given region begins to accumulate facial features building from skin tone through eyebrows/eyes and then to a mouth. The more facial features found for a given region which is a face candidate, the greater the possibility that the candidate actually is a face.

Furthermore, the above described method has the advantage that it is able to
10 cater for the circumstance where a face is backgrounded against a background region of substantially the same colour. Under these circumstances, in Yang and Waibel's method, no boundary between the face and the background would be likely to be detected. Therefore the region as a whole would be selected for further testing. However, the above method utilises the full colour space to segment the image, before making
15 decisions about which pixels are skin colour. Consequently the face is more likely to be separated from the background. In addition, the method is naturally independent of orientation or partial occlusion of the face.

Furthermore, the above method also permits false positives to be examined at the further stage and therefore does not exclude from subsequent testing regions which are
20 likely to be ultimately determined as a facial region.

The first embodiment described above notes that the nature of the original image may be taken into account when performing an initial face detection process. Further embodiments to be now described build upon this feature.

When an image is captured using a camera it is necessary either for the person
25 taking the picture to manually establish the camera settings (such as shutter speed,

aperture, focal length, etc), or for the camera to perform this operation automatically. Whichever is the case, the settings of the camera directly effect the appearance and quality of the image taken. In particular, the perceived brightness, colour, and sharpness of the objects within an image all depend on how the settings of the camera are
5 configured. For example, it is possible to take two pictures of the same scene with different camera settings and to obtain two images in which the same objects appear with different colours and brightness. Therefore, the ability to calibrate (in particular) colour information contained in (digital) images enables a broad variety of object detection and classification tasks in which colour is a strong discriminating feature.

- 10 Face detection is one such example application, and the present inventors have determined that the creation of face colour distribution models (CDM's), each adapted to specific lighting conditions, that can improve both the accuracy and reliability of face detection. Variations in lighting conditions can result from the use of a flash, such being a feature recognised as contributing in the face detection method of the first embodiment.
- 15 Since lightness is representative of colour features such as luminance and chrominance, such features may be used to quantify face detection.

- Before an image can be processed using a face colour distribution model, the face colour distribution model must be constructed. This is performed according to a method 50 shown in Fig. 5. The method 50 firstly gathers image samples at step 52 that
20 are representative images that contain faces, the images being acquired under a variety of lighting conditions and thus indicative of changes in luminance and chrominance. These images are then manually examined in step 54 to extract regions of skin colour for further processing in model formation. Step 54 may be performed by manually drawing a bounding box around a sample of face coloured pixels. Step 56, which follows, derives
25 colour representation values for the extracted pixels. This may be performed by

transforming the extracted pixels into a perceptual colour space such as CIE L^*u^*v or CIE L^*a^*b , so that each pixel is represented by at least a 2-dimensional vector. Alternatively other colour spaces such as HLS and HSV may be used. Preferably each pixel is represented as a length-3 vector incorporating both the luminance and chrominance values.

The colour representation values of pixels are then divided at step 58 into a number of sets (58a, 58b ... 58n) according to the lighting conditions present when each of the images were captured. Example sets are flash, non-flash, indoor, outdoor, and combinations of these. Alternatively, lighting parameters obtained directly from the camera such as the operation of a flash, may be used to identify and distinguish the sets. Other lighting conditions such as bright or cloudy, dusk or dawn, or a type of artificial light such as fluorescent, incandescent or halogen, may be used or detected for these purposes. These details may be provided by means of human input at the time of image capture.

For each of the sets (58a ... 58n) of face samples, step 60 then constructs a corresponding colour distribution model (CDM) (60a ... 60n) that best fits the samples of face colour pixels. The CDM can be a histogram, a probability density function, or a binary map. In one embodiment, a mixture of Gaussian PDF's are fit to the sample data using techniques known in the art such as the expectation maximisation (EM) algorithm, with either cross-validation, jackknife, and bootstrap techniques being used to estimate the goodness of fit of the model.

When each CDM (60a ... 60n) has been constructed, it is then desirable as shown in step 62 to establish a corresponding probability threshold (62a ... 62n) below which a colour vector is to be classified as relating to a non-face pixel, and above which the colour vector is to be classified as a potential face pixel. Additionally, the face colour

probability can be used directly in further facial image analysis steps detailed below. In the preferred embodiment, the CDM is constructed from colour representation values derived using a perceptual colour space (such as CIE L^*u^*v or CIE L^*a^*b) and then transformed back into the colour format of the input image, ie., either RGB or YUV. This
5 removes the necessity for transforming the input image into the perceptual colour space.

Since different image capture devices have differing performance, often determined by the quality and size of optical components (eg. lens, mirrors, aperture etc.), typically a CDM or a set of CDM's are generated for a particular capture device. In one implementation, where the image capture device (eg. camera) includes a light meter, a
10 reading from the light meter at the moment the image was captured can be used to determine the required CDM. In this fashion, a greater range of colour models may be devised and can be selected without possible human interference. Such interference may occur where the human user manually selects the operation of the flash where otherwise automatic operation of the flash would not be required. Further, the flash/outdoors
15 example sets above give rise to four (4) sets of CDM's. Using a light meter with, say, 4-bit encoding, can provide sixteen (16) models. Also, use of a light meter provides for enhanced reproducibility of results and enables the face samples used to generate the models to be taken under laboratory conditions and installation at the time of camera manufacture.

20 The processing 70 of an image according to a second embodiment is shown in Fig. 6. An input image is provided at step 72 and at step 74 the lighting conditions under which the image was captured are determined. Such a determination may be based on binary data obtained directly from the camera (eg. flash+indoors, no_flash+outdoors, no_flash+indoors, flash+outdoors) or corresponding meta-data provided with, or
25 accompanying, the image, which may be encoded or otherwise communicated according

to a predetermined format. Once the lighting conditions are determined, a corresponding or closest CDM is selected from a bank of look-up tables 78 retaining the CDM's (60a ... 60n) previously determined. At step 80, a first pixel of the input image 72 is selected and at step 82 is tested to see if the (RGB or YUV) colour of the pixel is contained in the selected CDM (60a ... 60n).

The steps shown in Fig. 6 following the comparison of step 82 depend upon the manner in which the CDM's are stored. In a preferred implementation, the selected thresholds of step 62 (Fig. 5) are used to construct a binary map or look-up table, where a representative colour vector is represented by a 1 if it is a colour vector that is contained within the thresholded face colour distribution, and by a 0 if the colour vector does not occur in the thresholded colour distribution. Alternatively, the CDM may represent the frequencies of the representative colour vectors of the thresholded colour distribution (ie. the CDM is effectively a histogram of representative colour vectors). A further variation is the case where a sampled distribution is approximated by a parametric model such as a Gaussian or a mixture of Gaussians. In the latter case, the CDM comprises the parameters of the model (eg. mean, co-variance)

As seen in Fig. 6, and according to the preferred implementation, a 1 or 0 value arising from step 82 is added to a map in step 84. Step 86 determines if there are more pixels in the image to be processed and step 88 obtains and passes the next pixel to step 82 for appropriate testing. When all pixels have been tested against the selected CDM, step 90 indicates the result of the preceding steps as being a binary face image map formed using detected skin-coloured pixels.

The map is then subjected at step 92 to further analysis of the skin-coloured pixels to provide at step 94 a face detection map for the image. The further analysis of

step 92 is, like the first embodiment, preferably independent of considerations of facial colour.

In practice, binary face map formed at step 90 may contain areas where either there are small non-face pixels (0's) surrounded by face pixels (1's), or vice versa. One approach for further analysis according to step 92 is processing the binary face image map so as to set to 0 any pixel locations which are contained in areas that are smaller than the smallest size of a potential face and to set any 0 pixel locations to 1 if they are surrounded by likely face colour pixels. This may be performed using a pair of morphological opening and closing operations with suitably shaped structuring elements.

- 10 A first structuring element such as:

$$\begin{bmatrix} 0 & 1 & 1 & 0 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 0 \end{bmatrix},$$

is used in the opening operation to remove potential face candidate pixel locations below this size. A second structuring element such as:

$$\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix},$$

- 15 is used in the closing operation to fill any holes in potential face candidate pixel locations.

Alternative approaches to the use of the structuring elements include using a Hough transform, or to count the number of pixels in the region having skin colour and to threshold that count against a predetermined percentage value. Other methods may be used to perform these tasks.

- 20 The result of the process 70 of Fig. 6 is a face detection map of pixel locations in the input image at which a face has been detected, and in all likelihood, a face is present.

The aforementioned edge detection method of further processing the likely face pixels to determine if a face exists may then be performed on the face detection map 94 resulting from the method 70.

In the preferred embodiment, the face colour distribution models are built for a number of separate lighting conditions, such as flash, non-flash, indoor, outdoor, and the like. However, this technique may be extended to the more general case of arbitrary lighting conditions based directly on parameters obtained from the camera. A list of camera parameters that may be of use in this situation is as follows:

- (i) white balance;
- 10 (ii) white balance mode
- (iii) aperture (iris);
- (iv) shutter speed;
- (v) auto gain control (AGC);
- (vi) auto exposure (AE) mode;
- 15 (vii) gamma;
- (viii) pedestal level; and
- (ix) flare compensation.

The parameters obtained from the camera are preferably obtained from a meta-data stream associated with the capture of each image (or video sequence). Examples, of such transmission protocols include IEEE 1394 ("firewire"). Also the ISO standards have defined methods for attaching meta-data to images and video in MPEG-7, MPEG4, and JPEG.

Whilst the first embodiment described with reference to Figs. 1 to 3 divides the image according to regions of substantially homogeneous colour, the second embodiment, and a third embodiment, are not so constrained.

The third embodiment is depicted in Fig. 7 by a method 150 where an input image 152 is provided and processed according to steps 154, 156 and 158 corresponding to steps 74, 76 and 78, respectively, of the second embodiment. Once the appropriate CDM has been selected in step 156, step 160 follows to process the input image as one or more regions. As a single region, the entirety of the image is processed on a pixel-by-pixel basis. Alternatively, the input image may be geometrically divided into simple pixel blocks (eg. 25 x 25 pixels, 10 x 20 pixels) which can be formed and processed in raster order. A further alternative, is where the regions, like the first embodiment, are divided on the basis of substantially homogeneous colour.

- 10 Step 162 selects a first region to be processed and step 164 a first pixel of that region. Step 166 compares the selected pixel with the CDM in a manner corresponding to step 82 of the second embodiment. Where the pixel matches the model, step 168 increments a count of pixels in the region meeting that criteria. Step 170 determines whether there are any other pixels in the region to be processed and, if so, step 172 obtains the next pixel and returns to step 166 for appropriate testing. When all pixels in the region have been processed, step 174 follows to compare the percentage of pixels classified for the region as skin colour against a predetermined percentage threshold value. Where the percentage is less than the predetermined number, the region is considered a non-face region and step 176 follows to test if there are any more regions to be processed. If so, step 178 selects the next region and returns processing to step 164. The count is then re-set. If not, the method 150 ends at step 184.

- Where the percentage exceeds the predetermined percentage, the region is considered a possible face region and step 180 follows to assess the region according to further facial detection analysis. Where such analysis does not detect a face, the method 150 proceeds to step 176 to process any further regions. Where the further

analysis of step 180 detects a face, step 182 registers that region as a face region and returns to step 176.

An example of the further analysis able to be performed as a consequence of appropriate marking at step 180 is the edge detection method described above in relation to the first embodiment.

The above described embodiments each indicate that face detection in images may be performed as a two stage process, a first representing something akin to a first filtering of the image to obtain likely candidate pixels or regions, and the second representing more thorough analysis to provide an actual determination on those pixels or regions passed by the first stage. In each case, lighting conditions associated with the capture of the image contribute to the determination performed by the first stage.

The above described methods are preferably practiced using a conventional general-purpose computer system 100, such as that shown in Fig. 4 where the processes of Fig. 3 and/or Figs. 5 and 6 are implemented as software, such as an application program executing within the computer system 100. In particular, the steps of the methods are effected by instructions in the software that are carried out by the computer. The software may be divided into two separate parts; one part for carrying out the above method steps; and another part to manage the user interface between the latter and the user. The software may be stored in a computer readable medium, including the storage devices described below, for example. The software is loaded into the computer from the computer readable medium, and then executed by the computer. A computer readable medium having such software or computer program recorded on it is a computer program product. The use of the computer program product in the computer preferably effects an advantageous apparatus for detecting face candidate regions in accordance with the embodiments of the invention.

The computer system 100 comprises a computer module 101, input devices such as a keyboard 102 and mouse 103, output devices including a printer 115 and a display device 114. A Modulator-Demodulator (Modem) transceiver device 116 is used by the computer module 101 for communicating to and from a communications network 120, for example connectable via a telephone line 121 or other functional medium. The modem 116 can be used to obtain access to the Internet, and other network systems, such as a Local Area Network (LAN) or a Wide Area Network (WAN), these being possible sources of input images and destinations for detected faces.

The computer module 101 typically includes at least one processor unit 105, a memory unit 106, for example formed from semiconductor random access memory (RAM) and read only memory (ROM), input/output (I/O) interfaces including a video interface 107, and an I/O interface 113 for the keyboard 102 and mouse 103 and optionally a joystick (not illustrated), and an interface 108 for the modem 116. A storage device 109 is provided and typically includes a hard disk drive 110 and a floppy disk drive 111. A magnetic tape drive (not illustrated) may also be used. A CD-ROM drive 112 is typically provided as a non-volatile source of data. The components 105 to 113 of the computer module 101, typically communicate via an interconnected bus 104 and in a manner which results in a conventional mode of operation of the computer system 100 known to those in the relevant art. Examples of computers on which the embodiments can be practised include IBM-PC's and compatibles, Sun Sparcstations or alike computer systems evolved therefrom.

Typically, the application program of the preferred embodiment is resident on the hard disk drive 110 and read and controlled in its execution by the processor 105. Intermediate storage of the program and any data fetched from the network 120 may be accomplished using the semiconductor memory 106, possibly in concert with the hard

disk drive 110. In some instances, the application program may be supplied to the user encoded on a CD-ROM or floppy disk and read via the corresponding drive 112 or 111, or alternatively may be read by the user from the network 120 via the modem device 116. Still further, the software can also be loaded into the computer system 100 from other
5 computer readable medium including magnetic tape, a ROM or integrated circuit, a magneto-optical disk, a radio or infra-red transmission channel between the computer module 101 and another device, a computer readable card such as a PCMCIA card, and the Internet and Intranets including e-mail transmissions and information recorded on Websites and the like. The foregoing is merely exemplary of relevant computer readable
10 mediums. Other computer readable mediums may be practiced without departing from the scope and spirit of the invention.

The further processing to candidate face images and regions may also be performed by or using the computer system 100 and known arrangements for such processing.

15 The method of detecting face candidate regions may alternatively be implemented in dedicated hardware such as one or more integrated circuits performing the functions or sub functions of Fig. 3 and/or Figs. 5 and 6. Such dedicated hardware may include graphic processors, digital signal processors, or one or more microprocessors and associated memories.

20

Industrial Applicability

It is apparent from the above that the embodiments of the invention are applicable in fields such as content-based image retrieval, personal identification or verification for use with automatic teller machines or security cameras, or automated interaction between humans and computational devices.

The foregoing describes only some embodiments of the present invention, and modifications and/or changes can be made thereto without departing from the scope and spirit of the invention as defined in the claims.

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
1029
1030
1031
1032
1033
1034
1035
1036
1037
1038
1039
1040
1041
1042
1043
1044
1045
1046
1047
1048
1049
1050
1051
1052
1053
1054
1055
1056
1057
1058
1059
1060
1061
1062
1063
1064
1065
1066
1067
1068
1069
1070
1071
1072
1073
1074
1075
1076
1077
1078
1079
1080
1081
1082
1083
1084
1085
1086
1087
1088
1089
1090
1091
1092
1093
1094
1095
1096
1097
1098
1099
1100
1101
1102
1103
1104
1105
1106
1107
1108
1109
1110
1111
1112
1113
1114
1115
1116
1117
1118
1119
1120
1121
1122
1123
1124
1125
1126
1127
1128
1129
1130
1131
1132
1133
1134
1135
1136
1137
1138
1139
1140
1141
1142
1143
1144
1145
1146
1147
1148
1149
1150
1151
1152
1153
1154
1155
1156
1157
1158
1159
1160
1161
1162
1163
1164
1165
1166
1167
1168
1169
1170
1171
1172
1173
1174
1175
1176
1177
1178
1179
1180
1181
1182
1183
1184
1185
1186
1187
1188
1189
1190
1191
1192
1193
1194
1195
1196
1197
1198
1199
1200
1201
1202
1203
1204
1205
1206
1207
1208
1209
1210
1211
1212
1213
1214
1215
1216
1217
1218
1219
1220
1221
1222
1223
1224
1225
1226
1227
1228
1229
1230
1231
1232
1233
1234
1235
1236
1237
1238
1239
1240
1241
1242
1243
1244
1245
1246
1247
1248
1249
1250
1251
1252
1253
1254
1255
1256
1257
1258
1259
1260
1261
1262
1263
1264
1265
1266
1267
1268
1269
1270
1271
1272
1273
1274
1275
1276
1277
1278
1279
1280
1281
1282
1283
1284
1285
1286
1287
1288
1289
1290
1291
1292
1293
1294
1295
1296
1297
1298
1299
1300
1301
1302
1303
1304
1305
1306
1307
1308
1309
1310
1311
1312
1313
1314
1315
1316
1317
1318
1319
1320
1321
1322
1323
1324
1325
1326
1327
1328
1329
1330
1331
1332
1333
1334
1335
1336
1337
1338
1339
1340
1341
1342
1343
1344
1345
1346
1347
1348
1349
1350
1351
1352
1353
1354
1355
1356
1357
1358
1359
1360
1361
1362
1363
1364
1365
1366
1367
1368
1369
1370
1371
1372
1373
1374
1375
1376
1377
1378
1379
1380
1381
1382
1383
1384
1385
1386
1387
1388
1389
1390
1391
1392
1393
1394
1395
1396
1397
1398
1399
1400
1401
1402
1403
1404
1405
1406
1407
1408
1409
1410
1411
1412
1413
1414
1415
1416
1417
1418
1419
1420
1421
1422
1423
1424
1425
1426
1427
1428
1429
1430
1431
1432
1433
1434
1435
1436
1437
1438
1439
1440
1441
1442
1443
1444
1445
1446
1447
1448
1449
1450
1451
1452
1453
1454
1455
1456
1457
1458
1459
1460
1461
1462
1463
1464
1465
1466
1467
1468
1469
1470
1471
1472
1473
1474
1475
1476
1477
1478
1479
1480
1481
1482
1483
1484
1485
1486
1487
1488
1489
1490
1491
1492
1493
1494
1495
1496
1497
1498
1499
1500
1501
1502
1503
1504
1505
1506
1507
1508
1509
1510
1511
1512
1513
1514
1515
1516
1517
1518
1519
1520
1521
1522
1523
1524
1525
1526
1527
1528
1529
1530
1531
1532
1533
1534
1535
1536
1537
1538
1539
1540
1541
1542
1543
1544
1545
1546
1547
1548
1549
1550
1551
1552
1553
1554
1555
1556
1557
1558
1559
1560
1561
1562
1563
1564
1565
1566
1567
1568
1569
1570
1571
1572
1573
1574
1575
1576
1577
1578
1579
1580
1581
1582
1583
1584
1585
1586
1587
1588
1589
1590
1591
1592
1593
1594
1595
1596
1597
1598
1599
1600
1601
1602
1603
1604
1605
1606
1607
1608
1609
1610
1611
1612
1613
1614
1615
1616
1617
1618
1619
1620
1621
1622
1623
1624
1625
1626
1627
1628
1629
1630
1631
1632
1633
1634
1635
1636
1637
1638
1639
1640
1641
1642
1643
1644
1645
1646
1647
1648
1649
1650
1651
1652
1653
1654
1655
1656
1657
1658
1659
1660
1661
1662
1663
1664
1665
1666
1667
1668
1669
1670
1671
1672
1673
1674
1675
1676
1677
1678
1679
1680
1681
1682
1683
1684
1685
1686
1687
1688
1689
1690
1691
1692
1693
1694
1695
1696
1697
1698
1699
1700
1701
1702
1703
1704
1705
1706
1707
1708
1709
1710
1711
1712
1713
1714
1715
1716
1717
1718
1719
1720
1721
1722
1723
1724
1725
1726
1727
1728
1729
1730
1731
1732
1733
1734
1735
1736
1737
1738
1739
1740
1741
1742
1743
1744
1745
1746
1747
1748
1749
1750
1751
1752
1753
1754
1755
1756
1757
1758
1759
1760
1761
1762
1763
1764
1765
1766
1767
1768
1769
1770
1771
1772
1773
1774
1775
1776
1777
1778
1779
1780
1781
1782
1783
1784
1785
1786
1787
1788
1789
1790
1791
1792
1793
1794
1795
1796
1797
1798
1799
1800
1801
1802
1803
1804
1805
1806
1807
1808
1809
1810
1811
1812
1813
1814
1815
1816
1817
1818
1819
1820
1821
1822
1823
1824
1825
1826
1827
1828
1829
1830
1831
1832
1833
1834
1835
1836
1837
1838
1839
1840
1841
1842
1843
1844
1845
1846
1847
1848
1849
1850
1851
1852
1853
1854
1855
1856
1857
1858
1859
1860
1861
1862
1863
1864
1865
1866
1867
1868
1869
1870
1871
1872
1873
1874
1875
1876
1877
1878
1879
1880
1881
1882
1883
1884
1885
1886
1887
1888
1889
1890
1891
1892
1893
1894
1895
1896
1897
1898
1899
1900
1901
1902
1903
1904
1905
1906
1907
1908
1909
1910
1911
1912
1913
1914
1915
1916
1917
1918
1919
1920
1921
1922
1923
1924
1925
1926
1927
1928
1929
1930
1931
1932
1933
1934
1935
1936
1937
1938
1939
1940
1941
1942
1943
1944
1945
1946
1947
1948
1949
1950
1951
1952
1953
1954
1955
1956
1957
1958
1959
1960
1961
1962
1963
1964
1965
1966
1967
1968
1969
1970
1971
1972
1973
1974
1975
1976
1977
1978
1979
1980
1981
1982
1983
1984
1985
1986
1987
1988
1989
1990
1991
1992
1993
1994
1995
1996
1997
1998
1999
2000
2001
2002
2003
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014
2015
2016
2017
2018
2019
2020
2021
2022
2023
2024
2025
2026
2027
2028
2029
2030
2031
2032
2033
2034
2035
2036
2037
2038
2039
2040
2041
2042
2043
2044
2045
2046
2047
2048
2049
2050
2051
2052
2053
2054
2055
2056
2057
2058
2059
2060
2061
2062
2063
2064
2065
2066
2067
2068
2069
2070
2071
2072
2073
2074
2075
2076
2077
2078
2079
2080
2081
2082
2083
2084
2085
2086
2087
2088
2089
2090
2091
2092
2093
2094
2095
2096
2097
2098
2099
2100
2101
2102
2103
2104
2105
2106
2107
2108
2109
2110
2111
2112
2113
2114
2115
2116
2117
2118
2119
2120
2121
2122
2123
2124
2125
2126
2127
2128
2129
2130
2131
2132
2133
2134
2135
2136
2137
2138
2139
2140
2141
2142
2143
2144
2145
2146
2147
2148
2149
2150
2151
2152
2153
2154
2155
2156
2157
2158
2159
2160
2161
2162
2163
2164
2165
2166
2167
2168
2169
2170
2171
2172
2173
2174
2175
2176
2177
2178
2179
2180
2181
2182
2183
2184
2185
2186
2187
2188
2189
2190
2191
2192
2193
2194
2195
2196
2197
2198
2199
2200
2201
2202
2203
2204
2205
2206
2207
2208
2209
2210
2211
2212
2213
2214
2215
2216
2217
2218
2219
2220
2221
2222
2223
2224
2225
2226
2227

WE CLAIM:

1. A method of detecting a face in a colour digital image, said method comprising the steps of:
 - 5 (i) segmenting said image into a plurality of regions each having a substantially homogenous colour;
 - (ii) testing the colour of each said region created in step (i) to determine those regions having predominantly skin colour; and
 - (iii) 10 subjecting only the regions determined in step (ii) to further facial feature analysis whereby said regions created in step (i) not having a predominantly skin colour are not subjected to said further feature analysis.
2. The method as claimed in claim 1 including the further step of using in step (ii) a colour distribution model utilising previously sampled data.
- 15 3. The method as claimed in claim 1 including the further step of using chromatic colours derived from RGB values in step (ii).
4. The method as claimed in claim 1 wherein said image is from a camera.
- 20 5. The method as claimed in claim 4 wherein a particular image is selected from the group consisting of images taken with use of a flash, and images taken without use of a flash.

6. The method as claimed in claim 1 wherein the further facial analysis in step (iii) is independent of facial colour.

7. The method as claimed in claim 5 wherein the further facial analysis comprises
5 edge detection.

8. The method as claimed in claim 6 wherein said edge detection utilises an edge detection filter comprising second derivative Gaussian function in a first direction substantially orthogonal to the edge to be detected and a Gaussian function in a second
10 direction substantially orthogonal to said first direction.

9. The method as defined in claim 6 or 7 wherein further facial analysis is carried out utilising the spatial relationship of detected edges.

10. The method as defined in claim 1 wherein step (i) includes growing a region by determining the average colour of a group of pixels, determining the difference between the average colour of the group and the average colour of pixels adjacent to the group, and, if the colour difference is less than a predetermined threshold, adding the adjacent pixels to the group.
15

20

11. Apparatus for detecting a face in a colour digital image, said apparatus comprising,

segmenting means to segment said image into a plurality of regions each having a substantially homogeneous colour,

colour detecting means coupled to said segmenting means to determine those regions having predominantly skin colour,

analysis means coupled to said colour detection means to subject only those regions having predominantly skin colour to a facial feature analysis.

5

12. The apparatus as claimed in claim 11 wherein said colour detection means includes a colour distribution model and storage means storing previously sampled data for use in said colour distribution model.

10 13. The apparatus as claimed in claim 12 including chromatic colour calculations means to derive from RGB values of said image, chromatic colour values.

14. The apparatus as claimed in claim 11 wherein said image is from a camera.

15 15. The apparatus as claimed in claim 11 wherein the particular image is selected from the group consisting of images taken with use of a flash, and images taken without use of a flash.

20 16. The apparatus as claimed in claim 11 wherein said facial feature analysis of said analysis means is independent of facial colour.

17. The apparatus is claimed in claim 10 wherein said analysis means includes an edge detector to detect edges in the image.

}

18. The apparatus as claimed in claim 17 wherein said analysis means has an edge detection filter comprising second derivative Gaussian function in a first direction substantially orthogonal to the edge to be detected and a Gaussian function in a second direction substantially orthogonal to said first direction.

5

19. The apparatus as claimed in claim 17 wherein said analysis means includes a spatial determinator to determine the spatial relationship of detected edges.

20. The apparatus as claimed in claim 11 including an accumulator means and a colour difference means to determine the difference between the average colour of a group of pixels and the average colour of pixels adjacent to that group, since that if the colour distance is less than a predetermined threshold, said adjacent pixels are added to said accumulator means to thereby grow a region.

21. A computer readable medium incorporating a computer program product for detecting a face in a colour digital image, said computer program product including a sequence of computer implementable instructions for carrying out the steps of:

(i) segmenting said image into a plurality of regions each having a substantially homogenous colour;

(ii) testing the colour of each said region created in step (i) to determine those regions having predominantly skin colour; and

(iii) subjecting only the regions determined in step (ii) to further facial feature analysis whereby said regions created in step (i) not having a predominantly skin colour are not subjected to said further feature analysis.

25

22. The medium as claimed in claim 21 including the further step of using in step (ii) a colour distribution model utilising previously sampled data.

23. The medium as claimed in claim 21 including the further step of using chromatic
5 colours derived from RGB values in step (ii).

24. The medium as claimed in claim 21 wherein said image is from a camera.

25. The medium as claimed in claim 21 wherein the particular image is selected
10 from the group consisting of images taken with use of a flash, and images taken without use of a flash.

26. The method as claimed in claim 21 wherein the further facial analysis in step (iii) is independent of facial colour.

15

27. The medium as claimed in claim 26 wherein the further facial analysis comprises edge detection.

28. The medium as claimed in claim 27 wherein said edge detection utilises an edge
20 detection filter comprising second derivative Gaussian function in a first direction substantially orthogonal to the edge to be detected and a Gaussian function in a second direction substantially orthogonal to said first direction.

29. The medium as defined in claim 26 wherein further facial analysis is carried out
25 utilising the spatial relationship of detected edges.

30. The medium as defined in claim 21 wherein step (i) includes growing a region by determining the average colour of a group of pixels, determining the difference between the average colour of the group and the average colour of pixels adjacent to the group, and, if the colour difference is less than a predetermined threshold, adding the adjacent pixels to the group.

31. A method of detecting a face in a colour digital image formed of a plurality of pixels, said method comprising the steps of:

10 (i) testing the colour of said pixels to determine those said pixels having predominantly skin colour, said testing utilising at least one image capture condition provided with said image; and

(ii) subjecting only said those pixels determined in step (i) as having predominantly skin colour to further facial feature analysis whereby those said pixels not having a predominantly skin colour are not subjected to said further facial feature analysis.

32. A method according to claim 31 wherein each said image capture condition is acquired at a time said image is captured.

20

33. A method according to claim 32 wherein said image is encoded according to a predetermined format and said at least one image capture condition is represented as meta-data associated with said format.

34. A method according to claim 31 wherein said at least one image capture condition comprises lighting conditions at a time said image was captured.

35. A method according to claim 31 wherein step (i), comprises the sub-step,
5 preceding said testing, of:

(a) dividing said image into a plurality of regions, each said region comprising a plurality of said pixels; and

wherein said testing is performed on pixels within each said region to determine those ones of said regions that are predominantly skin colour and step (ii) comprises
10 performing said further facial feature analysis on only those said regions determined to be predominantly of skin colour.

36. A method according to claim 31 wherein step (i) utilises at least one predetermined colour distribution model, said model having been generated using
15 previously sampled facial image data.

37. A method according to claim 36 wherein said colour distribution model is generated for a particular image capture device.

20 38. A method according to claim 36 wherein separate colour distribution models are generated for said different image capture conditions.

39. A method according to claim 38 wherein said at least one image capture condition comprises lighting conditions at a time said image was captured and separate

colour models are generated for different lighting conditions at a time said previously sampled facial image data was captured.

40. A method according to claim 39 wherein separate colour distribution models are
5 generated for groups of images taken with a flash and images taken without a flash.

41. A method according to claim 39 wherein separate colour distribution models are generated for groups of images taken indoors and images taken outdoors.

10 42. A method according to claim 36 wherein each said colour distribution model is represented as a frequency histogram of colour representation vectors.

43. A method according to claim 36 wherein each said colour distribution model is represented as a probability distribution of colour representation vectors.

15 44. A method according to claim 36 wherein each said colour distribution model is represented as a binary map of colour representation vectors.

45. A method according to claim 42, 43 or 44 wherein said colour representation
20 vectors are derived from perceptual colour space values of the predetermined skin-colour pixels in said previously sampled facial image data.

46. A method according to claim 42, 43 or 44 wherein said colour representation
vectors contain chromatic colour values derived from those RGB values of the
25 predetermined skin-colour pixels in said previously sampled facial image data.

47. A method according to claim 44 wherein said binary map comprises a percentage of the skin colour pixels that were identified in said previously sampled facial image data.

5

48. A method according to claim 47 wherein one of said pixels is classified as being skin colour if the colour representation vector corresponding thereto occurs within said binary map.

10 49. A method according to claim 42 wherein each of said pixels is classified as being skin colour if the frequency of the colour representation vector corresponding thereto exceeds a predetermined threshold frequency.

15 50. A method according to claim 44 wherein each of said pixels is classified as being skin colour if the probability of the colour representation vector corresponding thereto exceeds a predetermined probability threshold.

20 51. A method according to claim 48 wherein one said regions is determined to be predominantly skin colour if more than a predetermined percentage of the total number of said pixels in said one region are classified as being skin colour.

52. A method according to claim 35 wherein said regions are geometrically divided from said image.

5

53. A method according to claim 35 wherein said regions are formed of pixels having substantially homogenous colour.

54. A method according to claim 53 wherein said regions are formed using a region
5 growing method based upon colour differences.

55. A method according to claim 35 wherein said further analysis of step (ii) is independent of face colour.

10 56. Apparatus for detecting a face in a colour digital image formed of a plurality of pixels, said apparatus comprising:

means for testing the colour of said pixels to determine those said pixels having predominantly skin colour, said testing utilising at least one image capture condition provided with said image; and

15 means for subjecting only said those pixels so determined as having predominantly skin colour to further facial feature analysis whereby those said pixels not having a predominantly skin colour are not subjected to said further facial feature analysis.

20 57. Apparatus according to claim 56 wherein each said image capture condition is acquired at a time said image is captured.

58. Apparatus according to claim 57 wherein said image is encoded according to a predetermined format and said at least one image capture condition is represented as
25 meta-data associated with said format.

59. Apparatus according to claim 56 wherein said at least one image capture condition comprises lighting conditions at a time said image was captured.

5 60. Apparatus according to claim 56 wherein said means for testing comprises means for dividing said image into a plurality of regions, each said region comprising a plurality of said pixels;

wherein said means for testing operates on pixels within each said region to determine those ones of said regions that are predominantly skin colour and said means
10 for subjecting cause said further facial feature analysis to be performed on only those said regions determined to be predominantly of skin colour.

61. A computer readable medium incorporating a computer program product for detecting a face in a colour digital image formed of a plurality of pixels, said computer
15 program product comprising:

means for testing the colour of said pixels to determine those said pixels having predominantly skin colour, said testing utilising at least one image capture condition provided with said image; and

means for subjecting only said those pixels so determined as having
20 predominantly skin colour to further facial feature analysis whereby those said pixels not having a predominantly skin colour are not subjected to said further facial feature analysis.

62. A computer readable medium according to claim 61 wherein each said image
25 capture condition is acquired at a time said image is captured.

63. A computer readable medium according to claim 62 wherein said image is encoded according to a predetermined format and said at least one image capture condition is represented as meta-data associated with said format.

5

64. A computer readable medium according to claim 61 wherein said at least one image capture condition comprises lighting conditions at a time said image was captured.

65. A computer readable medium according to claim 61 wherein said means for testing comprises means for dividing said image into a plurality of regions, each said region comprising a plurality of said pixels;

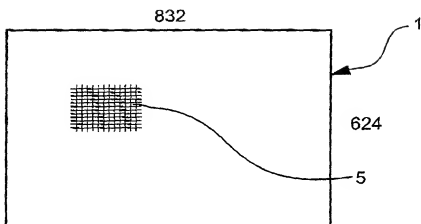
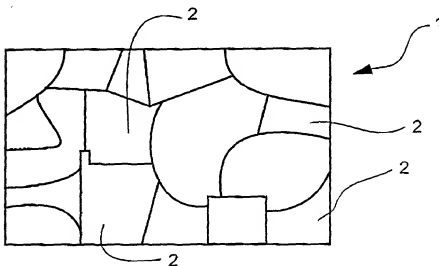
10 wherein said means for testing, operates on pixels within each said region to determine those ones of said regions that are predominantly skin colour and said means for subjecting cause said further facial feature analysis to be performed on only those said regions determined to be predominantly of skin colour.

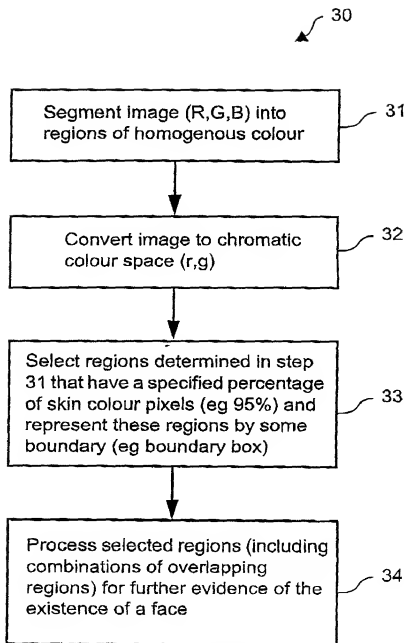
15

Abstract

FACE DETECTION IN DIGITAL IMAGES

- 5 The present invention relates to the detection of faces in digital images. Rather than
subjecting the entire image (1) to computationally intensive face detection analysis, the
image is subjected to computationally simple analysis to identify candidate pixels likely
to be of skin colour. Only those pixels having such colour, which may be determined in a
number of ways, are subject to computationally intensive face detection analysis. The
10 image is preferably divided into regions (2) each of which is then analysed for features
indicative of skin colour. The simple analysis is based on lighting conditions associated
with the capture of the image and preferably utilizes the selection of one of a number of
skin colour models associated with a range of lighting conditions.

**Fig. 1****Fig. 2**

**Fig. 3**

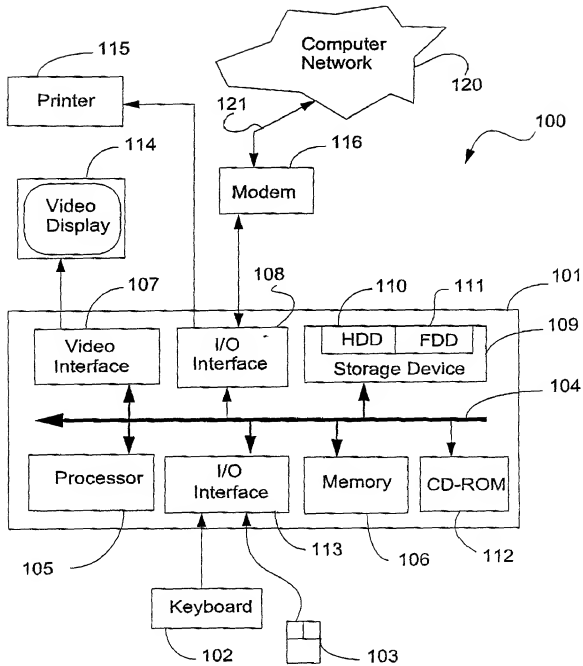
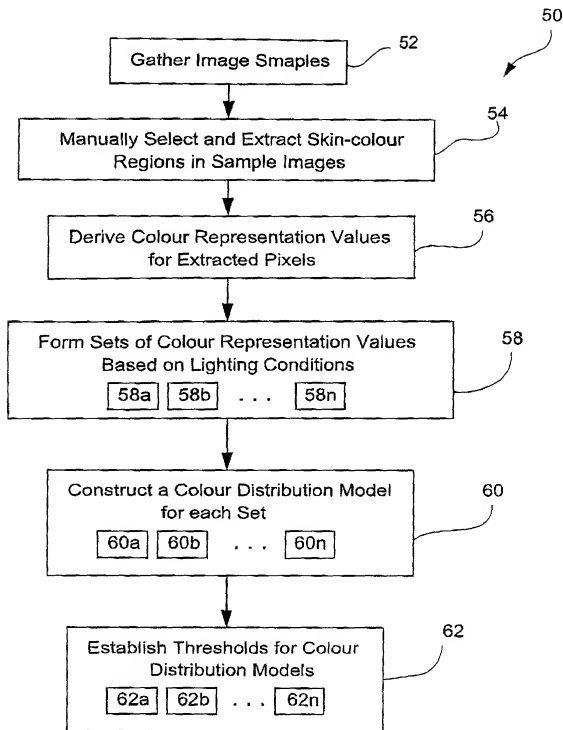


Fig. 4

**Fig. 5**

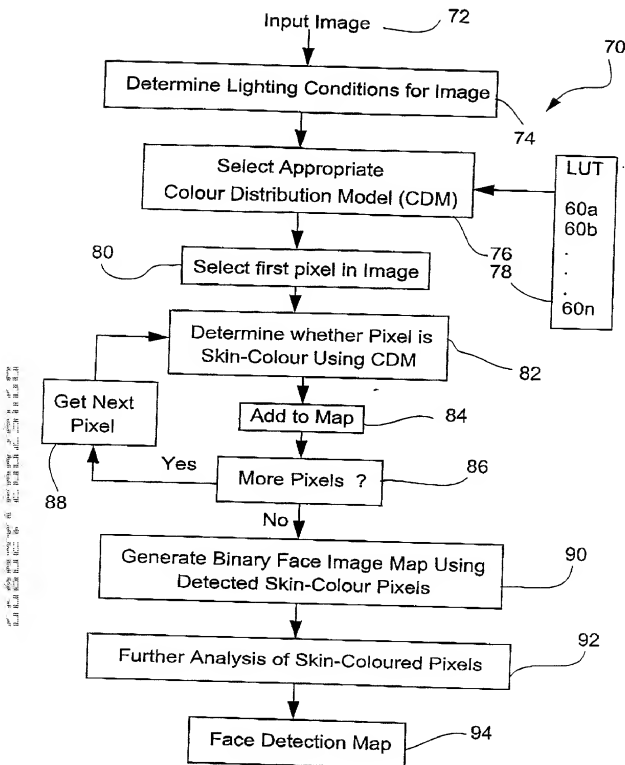


Fig. 6

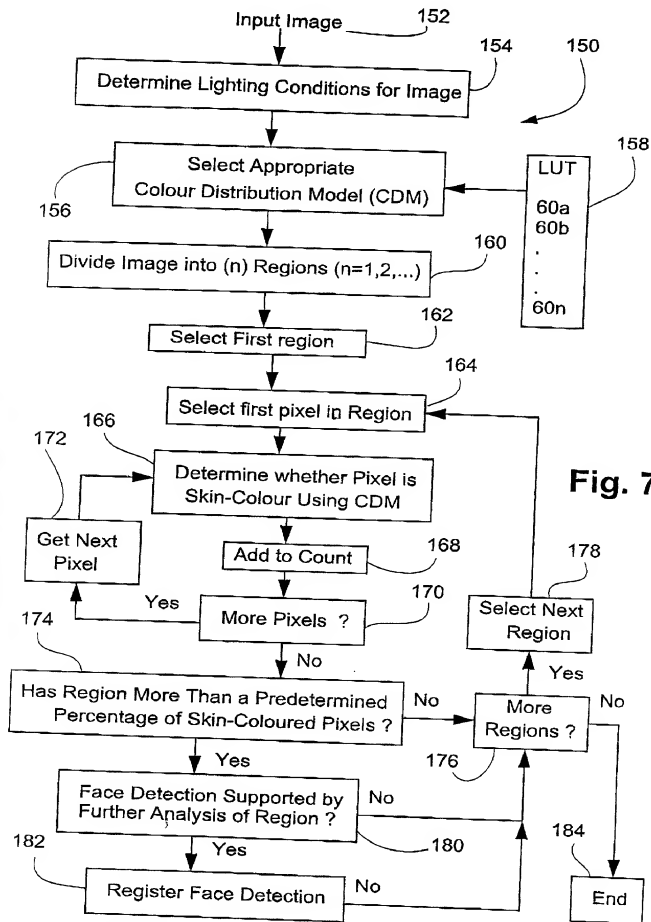


Fig. 7

**COMBINED DECLARATION AND POWER OF ATTORNEY
FOR C-I-P PATENT APPLICATION**

As a below named inventor, I hereby declare that.

My residence, post office address and citizenship are as stated below next to my name.

I believe I am the original, first and sole inventor (if only one name is listed below) or an original, first and joint inventor (if plural names are listed below) of the subject matter which is claimed and for which a patent is sought on the invention entitled FACE DETECTION IN DIGITAL IMAGES

_____, the specification of which is

☒ attached hereto: ☐ was filed on _____
as Application No. _____ and was amended on _____
(if applicable).

I hereby state that I have reviewed and understand the contents of the above-identified specification, including the claims, as amended by any amendment referred to above.

I acknowledge the duty to disclose information which is material to patentability as defined in 37 CFR §1.56.

I hereby claim foreign priority benefits under Title 35, United States Code, §119 of any foreign application(s) for patent or inventor's certificate listed below and have also identified below any foreign application for patent or inventor's certificate having a filing date before that of the application on which priority is claimed:

Country	Application No.	Filed (Day/Mo./Yr.)	(Yes/No) Priority Claimed
AUSTRALIA	PP4009	10 June 1998	YES

I hereby claim the benefit under Title 35, United States Code, §120 of any United States application(s) listed below and, insofar as the subject matter of each of the claims of this application is not disclosed in the prior United States application in the manner provided by the first paragraph of Title 35, United States Code §112, I acknowledge the duty to disclose material information as defined in Title 37, Code of Federal Regulations, §1.56(a) which occurred between the filing date of the prior application and the national or PCT international filing date of this application:

Application No.	Filed (Day/Mo./Yr.)	Status (Patented/Pending/Abandoned)
09/326,561	07 June 1999	Pending

I hereby appoint the practitioners associated with the firm and Customer Number provided below to prosecute this application and to transact all business in the Patent and Trademark Office connected therewith, and direct that all correspondence be addressed to the address associated with that Customer Number.

FITZPATRICK, CELLA, HARPER & SCINTO
Customer Number: 05514

I hereby declare that all statements made herein of my own knowledge are true and that all statements made on information and belief are believed to be true; and further that these statements were made with the knowledge that willful false statements and the like so made are punishable by fine or imprisonment, or both, under Section 1001 of Title 18 of the United States Code and that such willful false statements may jeopardize the validity of the application or any patent issued thereon.

**COMBINED DECLARATION AND POWER OF ATTORNEY
FOR C-I-P PATENT APPLICATION**
(Page 2)

Full Name of Sole or First Inventor EDWIN HO
Inventor's signature _____
Date _____ Citizen/Subject of AUSTRALIA
Residence 29 Forman Avenue, Glenwood, New South Wales 2768, Australia
Post Office Address c/o CANON KABUSHIKI KAISHA
30-2, Shimomaruko 3-chome, Ohta-ku, Tokyo, Japan

Full Name of Second Joint Inventor, if any ALISON JOAN LENNON
Second Inventor's signature _____
Date _____ Citizen/Subject of GREAT BRITAIN
Residence 554 Darling Street, Balmain, New South Wales 2041, Australia
Post Office Address c/o CANON KABUSHIKI KAISHA
30-2, Shimomaruko 3-chome, Ohta-ku, Tokyo, Japan

Full Name of Third Joint Inventor, if any ANDREW PETER BRADLEY
Third Inventor's signature _____
Date _____ Citizen/Subject of AUSTRALIA
Residence 68 The Bulwark, Castlecrag, New South Wales 2068 Australia
Post Office Address c/o CANON KABUSHIKI KAISHA
30-2, Shimomaruko 3-chome, Ohta-ku, Tokyo, Japan

Full Name of Fourth Joint Inventor, if any _____
Fourth Inventor's signature _____
Date _____ Citizen/Subject of _____
Residence _____
Post Office Address _____

Full Name of Fifth Joint Inventor, if any _____
Fifth Inventor's signature _____
Date _____ Citizen/Subject of _____
Residence _____
Post Office Address _____

Full Name of Sixth Joint Inventor, if any _____
Sixth Inventor's signature _____
Date _____ Citizen/Subject of _____
Residence _____
Post Office Address _____

/PAP